Is participatory social learning a performance driver for Chinese smallholder farmers?

Huanxiu Guo, Sébastien Marchand

CENTRE D'ETUDES
ET DE RECHERCHES
SUR LE DEVELOPPEMENT
INTERNATIONAL

SERIE ETUDES ET DOCUMENTS DU CERDI

Is participatory social learning a performance driver for Chinese smallholder farmers?

Huanxiu Guo and Sébastien Marchand

Études et Documents **n°18**

Octobre 2013

CERDI
65 BD. F. MITTERRAND
63000 CLERMONT FERRAND - FRANCE
TEL. 04 73 71 74 20
FAX 04 73 17 74 28
www.cerdi.org

Les auteurs / The authors

Huanxiu Guo

Doctorant / PhD Student

Clermont Université, Université d'Auvergne, CNRS, UMR 6587, CERDI, F-63009 Clermont Fd, France.

Email : huanxiu.guo@etu-udamail.fr

Sebastien Marchand

Maitre de Conférence en économie

Clermont Université, Université d'Auvergne, CNRS, UMR 6587, CERDI, F-63009 Clermont Fd, France.

Email : sebastien.marchand@udamail.fr

La série des Etudes et Documents du CERDI est consultable sur le site :

http://www.cerdi.org/ed

Directeur de la publication : Patrick Plane

Directeur de la rédaction : Catherine Araujo Bonjean

Responsable d'édition : Annie Cohade

ISSN : 2114 7957

Résumé / Abstract

This paper aims to test the effect of smallholder farmers' participatory social learning on their gain of performance in a village of southwest China. By exploring a panel structure survey data collected in the village, we identify the social learning effect using a Spatial Autoregressive (SAR) model. Particularly, we calculate the technical efficiency and environmental efficiency from a SFA model and use them as dependent variables of the model. Moreover, we investigate the social learning of different technologies, i.e., conventional and organic farming, by separating the estimations. Our identification results suggest that the effect of social learning is weak due to the technological heterogeneity in the general case, whilst it is significantly positive for organized organic farming. However, it appears that farmers learn to improve their economic performance (i.e., maximize yield) rather than environmental performance (i.e., minimize environmentally detrimental input). These results reveal a critical limitation of social learning, and demand more environmental orientation in the agricultural extension service, which is expected to guide smallholder farmers and foster their environmental performance for sustainable agricultural development.

Mots clés / Keywords : Smallholder farming, Social learning, Organic farming, Technical efficiency, Environmental efficiency, China.

Codes JEL / JEL codes : Q12, Q57, D83, D24, O53

Remerciements / Acknowledgements

We thank the participants of the AFSE 2013 annual conference in Aix-en-Provence for their useful comments. We would like to thank the NGO Partnerships for Community Development (PCD) and the Guangxi Maize Research Institute for their valuable technical assistance in the field work. We are grateful for the financial support of this work from the Foundation of *Université d'Auvergne* (UDA).

1 Introduction

Smallholder or family farming is the primary and most widespread form of agricultural production in developing countries. It is estimated that 500 million rural people in developing countries live on small farms (less than two hectares). The majority are undernourished people living in absolute poverty (Hazell et al., 2007). According to FAO's World Census of Agriculture, there are 193 million small farms in China, which represent more than 95 per cent of total farms (Swaminathan, 2013; Bélières et al., 2013). Given the special land tenure regime, the average farm size in China is under 0.5 hectare (Fan and Chan-Kang, 2005)[1].

This prevalent smallholder farming plays a critical role in the sustainable agricultural development. From a sociological perspective, smallholder farming ensures the social equity, poverty reduction and food security. It is essential for poor people with limited resources and their substantial livelihood depends on these small pieces of land (Hazell and Ramasamy, 1991; Greenland, 1997). From an economic perspective, smallholder farming may be more productive according to the studies supporting the inverse relationship between farm size and productivity (Sen, 1962; Feder, 1985; Heltberg, 1998; Raghbendra et al., 2000; Fan and Chan-Kang, 2005; Lipton, 2006). From an ecological perspective, smallholder farming is natural resource and bio–diversity conserving, which makes it more suitable and favorable for the development of environmentally sound and sustainable agricultural technologies for our green future (Altieri, 2002).

However, smallholder farming is facing challenges in the context of economic globalization and transition. For instance, along with the development of manufacturing sector, more and more smallholder farmers intend to move out of agriculture. In China, more than 150 million farmers have moved out to work in the city (Cai and Wang, 2008; NBSC, 2012). Meanwhile, accelerating urbanization and attractive investment opportunities have opened up in agriculture, leading to large–scale investments and competition for land, e.g., rubber plantations in Cambodia, palm oil production in Indonesia, real

[1] In China, the agricultural land is collectively owned. Under the Household Responsibility System (HRS), rural households have right to exploit arable land for a long period of 30 years. The size of land mainly depends on household size and composition.

estate exploitation in China. Moreover, the soaring price of productive inputs, deterioration of agro–environment and threats of climate change make smallholder farming more vulnerable.

In recent years, it is increasingly acknowledged that smallholder farming should be revived in the development of sustainable agriculture. It is argued that smallholder farmers who get involved in the management of social-ecological systems may learn and therefore enhance their adaptive capacity through their involvement in decision making processes (Folke et al., 2005; Fazey et al., 2007). This process is known as social learning. In the influential work of Bandura (1977), social learning is defined as individual learning based on observation of others and their social interactions within a group, e.g. through imitation of role models. Supported by the social learning theory, a number of farmer participatory development approaches have been put forward to get smallholder farmers involved in the sustainable agriculture in developing countries (Pretty, 1995; Desai and Potter, 2013). The raising literature of social learning and participatory development has critical implication for ongoing sustainable agriculture development, yet the empirical test of social learning remains rare. To our knowledge, few empirical study of the social learning exists in China.

To fill in this blank, this paper attempts to determine the effect of participatory social learning on smallholder farmers' economic and environmental performances in rural China. We have identified the Sancha village in southwest China where farmer participatory social learning was organized for organic paddy rice production. We conduct a household survey in the village and collect a plot level panel dataset for the empirical analysis. In terms of econometric methodology, we combine the Spatial Autoregressive (SAR) model with peer effect analysis to estimate the social learning effect within carefully defined learning group. Specifically, we purge confounding factors such as inputs contamination (i.e., nitrogen fertilizer) by using the technical efficiency and environmental efficiency as dependent variables of the model. With different efficiency terms, we test whether farmers learn to maximize their output (technical efficiency), or to minimize their nitrogen use (environmental efficiency). Finally, the estimation is applied within separated sub–samples of conventional and organic farming to test the technological constraints of social learning.

Our estimation results suggest that the effect of social learning is non significant among smallholder farmers in general case. This is mainly due to the heterogeneity of technology in smallholder system. In the case of organized organic farming, social learning is significant in improving farmers' technical efficiency, but not their environmental efficiency. In other words, farmers learn to maximize their output rather than to minimize nitrogen input. Based on these results, we conclude that social learning is effective in fostering smallholder farmers' performance for productive agriculture if it is well organized. However, for the goal of resource conservative agriculture, external supports such as extension service and environmental education are needed to guide smallholder farmers.

For the remainder of the paper, Section 2 reviews the literature of social learning and smallholder sustainable agriculture. Section 3 describes the organization of social learning in the village. Section 4 explains the methodological framework of our analysis and the identification strategy. Section 5 gives details about our data and Section 6 discusses the main results. Section 7 provides policy implications and concludes.

2 A literature review of social learning and sustainable agricultural development

Environmentally–sound or ecological agriculture (e.g. organic farming, low-input agriculture and permaculture) has been promoted by governments and development agencies for sake of sustainable agricultural development during past decades (FAO, 2002; IFAD, 2002; WorldBank, 2009; Twarog, 2006). In contrast to conventional agriculture, ecological agriculture can generate outstanding environmental benefits and ecosystem services, e.g. reduction of soil erosion and pollution, improvement of soil fertility and bio diversity, and alleviation of dependence on chemical inputs (Swinton et al., 2007). Particularly, sustainable agriculture seems to be more profitable in developing countries given its features of low external–input and increasing yield albeit original low level (Stoop et al., 2002; Pretty et al., 2003).

A big challenge for development of the sustainable agriculture in developing countries remains to involve smallholder farmers. In response, new development initiatives such as Participatory Research and Development (PRD) and the Farmer Field Schools (FFS) have emerged to promote sustainable agriculture to smallholder farmers in developing countries (Braun et al., 2000; Godtland et al., 2004; Feder et al., 2004). These initiatives aim to introduce and adapt sustainable agriculture to local conditions by farmers' participatory field trial and then to diffuse successful experience through a social learning process (Pretty, 1995; Pretty and Uphoff, 2002). A growing body of studies have emerged recently to evaluate the impact of farmer participatory initiatives and understand the process of social learning (Godtland et al., 2004; Feder et al., 2004; Van den Berg and Jiggins, 2007).

In economics, the literature of social network analysis has opened a new perspective for more thorough understanding about farmers' social learning process (Romer, 1986; Lucas, 1988; Rogers, 1995). On the theory side, Besley and Case (1994) develop a dynamic model of learning to study farmers' adoption decision of new technology. In this model, the uncertainty about the profitability is a major concern for farmers' adoption of new technology. In a Bayesian learning process, farmers can learn from their own experience as well as others' behavior about the true profitability and update their own behavior. In a repeated equilibrium, interaction between farmers is necessary provided the information is a public good. Using this model, one can predict the diffusion path of the new technology.

Alternatively, Foster and Rosenzweig (1995) adapt a "target-input" model to explain farmers' learning about the optimal use of inputs with a new technology(Wilson, 1975; Jovanovic and Nyarko, 1994). In the setting of "target-input" model, the profitability of new technology is increasing with the accumulation of knowledge observed from neighbors. The accumulated knowledge allows farmers to learn about the target input rate or the optimal input level to merit adoption. Therefore, they argue that learning about input productivity is as important as profitability in the diffusion of new technology, while identification of social learning using information of input productivity or its rewards is more accurate than adoption behavior.

The "target-input" model is useful to explain the social learning based on "rule-of-thumb" learning behavior (Conley and Christopher, 2001)[2]. In situations where agent

[2]The rule-of-thumb learning rules can be defined as follows: at each period, each agent constructs his

cannot observe his neighbors' experience perfectly, or where neighbors' performance is essentially determined by unobserved individual characteristics and conditions, social learning may be weak. In other words, agents learn from similar neighbors only (Ellison and Fudenberg, 1993). In a more recent work, Young (2009) revises the existent models of diffusion and makes a comprehensive comparison of the social learning model with the contagion model in marketing (Mahajan et al., 1990) and the social influence model in sociology (Granovetter, 1978).

In spite of the rich implications of social learning theory, the empirical identification of social learning effect is not easy. In his seminar work on social interaction, Manski (1993) uses the term of "reflection problem" to describe the difficulty of disentangling endogenous social effect (e.g. social learning) from exogenous social effect (contextual effect) in a *linear-in-means* model[3]. Subsequently by the discussions of Brock and Durlauf (2001) and Moffitt (2001), social learning effect is often confounded with common environment conditions or other group correlations that do not necessarily entail social learning. It raises even more concerns in the agricultural context because agricultural production generally depends on the common growing conditions.

To achieve convincing identification of social learning, one condition is to well define the reference group within which the learning process takes place. Then, different strategies can be employed to identify the social learning effect. Among others, Munshi (2004) compares the social learning effect on adoption of different HYV crops (wheat and rice) in the same district and finds that social learning is weak in a heterogeneous population. Bandiera and Rasul (2006) assume the effect of correlated unobservable is monotonic and test for non-linearities predicted by a model of strategic interactions in social learning. By doing so, they find an inverse–U shape social learning effect which depends on the number of adopters in the social network in Mozambique. Conley and Udry (2010) exploit the timing of news about neighbors productivity to test for social learning effect in the fertilizer use in Pineapple production in Ghana. They take special care in construction of reference group and control for environmental factors and find a positive social learning.

posterior as a weighted average of his prior, his signal and the information he receives from neighbors.

[3]In the linear-in-means model, the outcome of each agent depends linearly on his own characteristics, on its mean characteristics and on the mean outcome of his reference group.

However, most of empirical studies investigate the effect of social learning on the adoption of new technology. Few has tested the social learning on the performance of new technology (except for the study of (Conley and Udry, 2010)). To the best of our knowledge, there is still no empirical evidence about social learning on the performance of sustainable agriculture in developing countries. To fill in this blank, we follow the literature of social network analysis and attempt to test the social learning effect on smallholder farmers' performance in the context of NRR in rural China. We will now turn to this special case and discuss the organization of social learning in the Sancha village.

3 The participatory social learning in Sancha village

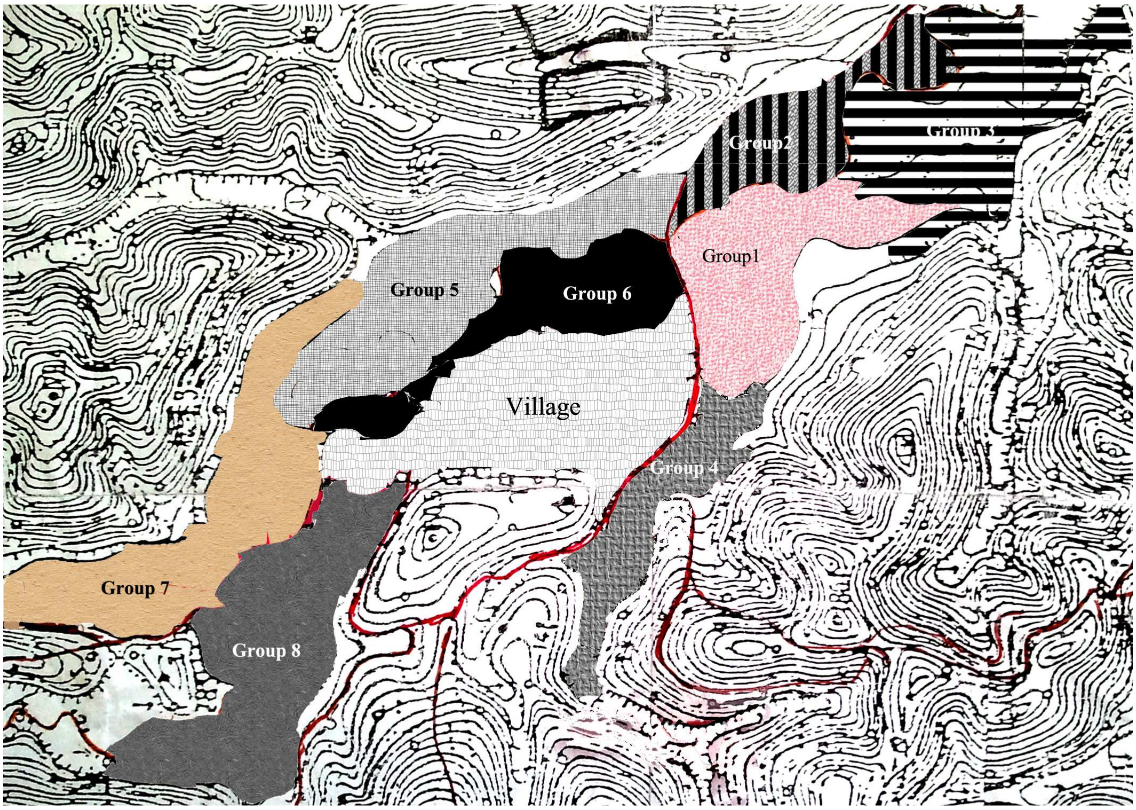
Sancha (109.01E/22.73N) is a small natural village located in the mountainous zone of Guangxi Zhuang Autonomous region[4]. Thanks to the well preserved environment, Sancha village was identified by a Hong Kong-based NGO, called Partnerships for Community Development (PCD), for a project of organic paddy rice production within the framework of "New Rural Reconstruction" in 2005. At the beginning, the organic rice production was introduced to a small group of farmers in the form of field trials. The PCD, in collaboration with the Guangxi Maize Research Institute (GMRI), had assisted farmers' experiments of organic farming with technical guidance and marketing support. Through these field trials, a number of local knowledge such as pest control with local medicinal plants, composting and the "Duck-Rice" integrated system had been revived to adapt organic farming to local conditions.

In order to diffuse the successful experience of organic farming and get more farmers involved, the project opted for an approach of participatory social learning. According to the investigation of PCD, the rice production was organized on basis of four families (i.e. Li, Xu, Huang and Lu families) in Sancha village. The paddy fields were thus divided and exploited by four families labelled production group 1, 2, 3 and 4. Specifically, each family

[4] A "village" in China can either be a natural village (Zirancun), one that spontaneously and naturally exists, or an administrative village (Xingzheng cun), which is a bureaucratic entity

has 2 groups of paddy fields, one in the plain zone and another in the mountainous zone. As such, 8 groups of paddy fields were naturally defined. Map 1 provides an overview for the location of these groups. Nowadays under the Household Responsibility System (HRS), collective farming system has been broken into individual production, whereas the grouping of paddy fields remains unchanged.

Map 1. Groups of paddy fields in Sancha village



Source: Sancha village committee

The participatory social learning took place within these 8 groups for several reasons. Firstly, farmers worked simultaneously in the group as paddy rice production is a seasonal activity. They communicated during agricultural work to exchange information and ease the work. Also, their coordination was frequent in the group due to their collective management of water resource. Secondly, provided the small size of paddy field, farmers could directly observe the inputs and outputs of other farmers in the group, which favored the learning process. Thirdly, farmers spoke the same language in the group, which was essential for the social learning. With this understanding, workshops were organized within the group where all farmers of the group were united together to share their experience of organic farming. Farmers could then learn from their peers to improve their management of the technology or start their own experiment of organic farming. The learning is qualified as a social process as it not only takes place within the workshops, but also extends to farmers' interaction during their daily production. It is expected that in a process of social learning, farmers can generate their own knowledge and reduce the dependance on external assistance.

In this study, we follow the PCD definition of learning group for an investigation of social learning effect. Farmers' learning peers are defined as all farmers of the same group except for himself. We recognize the limitation of such a rough measure of social learning networks, for it cannot distinguish the specific links between agents (e.g. relatives and friends). However, the definition takes account of all potential learning sources and avoids measurement errors and omission of information in the learning process, which is appropriate in a smallholder environment. With this definition of reference group, the social learning effect could be captured by the peer effect of the group outcome on the individual outcome (i.e. agricultural performance defined as yield, technical efficiency or environmental efficiency). Still, the social learning effect can be confounded with other environmental factors. For instance, the inputs contamination or the agro–ecosystem could also collectively affect farmers' agricultural performance. Moreover, farmers' agricultural performance could depend on other social mechanisms such as altruism that do not necessarily entail social learning. In the next section, we will present our identification strategy to disentangle social learning from these confounding factors.

4 Identification of the social learning effect

As discussed in the literature review, estimating social learning effect raises three main challenges. Firstly, Manski (1993) divided social effects into an endogenous part (i.e. social learning effect) and an exogenous part (i.e., contextual influence). The separation of these two parts is a main challenge and thoroughly discussed as the *reflection problem*[5]. Secondly, the problem of "correlated effects" will raise when social learning needs to be identified from confounding environmental effects (i.e. factors related to common group environment). Thirdly, spurious correlation among members of the same reference group plague the identification if the formation of group is endogenous(Moffitt, 2001)[6]. These challenges call for appropriate statistical methodology which differs from one study to another. We will follow the literature to present our strategy addressing these challenges in our specific setting.

4.1 The basic model

Our model is an extension of the standard linear-in-means social interaction model in which we allow for plot–specific reference groups. Consider we have a set of plots i, $(i = 1, \dots n)$. Let $y_{i,t}$ be the agricultural outcome (e.g. yield) of plot i in season t. Let $X_{i,t}$ be a vector of plot owner's characteristics. Each plot i may have a reference group P_i of size n_i. This reference group is known by the modeller and contains all plots whose outcomes or owner characteristics may affect plot i's outcome. The collection of plot–specific reference groups thus defines an undirected network of plots[7].

Consider the following spatial autoregressive regression (SAR) model extended to social network economics (Brock and Durlauf, 2001) in which spatial units are plots and the peers effects are either contextual or endogenous. Formally, the structural model is formulated as:

$$y_{i,t} = \alpha + \beta \frac{\sum_{j \in P_i} y_{j,t-1}}{n_i} + \theta y_{i,t-1} + \gamma X_{i,t} + \delta \frac{\sum_{j \in P_i} X_{j,t}}{n_i} + \tau_t + \epsilon_{i,t}. \qquad (1)$$

Note here we use the lagged peers' performance to avoid a potential simultaneity of outcome and β captures the endogenous social effect (i.e. social learning effect)[8]. δ captures

[5]Manski (1993) has pointed out that the expected outcome from social equilibrium might be linearly depended on observed exogenous variables of a group in the model, i.e., the *reflection problem*.

[6]Here is the case where members in a group share common characteristics which leads to self–selection for the creation of the group.

[7]This corresponds to usual empirical formulation (e.g. Lee (2007)).

[8]Using the lagged performance of neighbors avoids the simultaneity problems since the current performance of the plot i cannot explain the past performance of his peers. However, some correlated variables

the contextual effect and γ captures the effects of plot owner's characteristics (i.e. age, gender and education level). Given the underlying social process, the lagged errors $\epsilon_{i,t-1}$ are correlated with the lagged peers' outcomes and if the errors are serially correlated, the estimation of β is biased since the $cov(\epsilon_{i,t}, \frac{\sum_{j\in P_i} y_{j,t-1}}{n_i}) \neq 0$. The addition of lagged individual outcome $y_{i,t-1}$ thus helps to control for this potential bias. The τ_t is a dummy fixing five crop seasons in our data and the error term $\epsilon_{i,t}$ reflects the usual i.i.d. disturbances with zero mean and an unknown variance associated with i.

To make it simpler, we can write the structural model in matrix notation.

$$y_{i,t} = \alpha\iota + \beta G y_{j,t-1} + \theta y_{i,t-1} + \gamma X_{i,t} + \delta G X_{j,t} + \tau_t + \epsilon_{i,t}, \tag{2}$$

where y is an $n \times 1$ vector of outcomes, G is an $n \times n$ interaction matrix with $G_{ij} = 1/n_i$ if j and i are in the same reference group, and 0 otherwise, and ι is an $n \times 1$ vector of ones. G derives from our definition of reference group on basis of the village's social and geographical structure: (1) four families and (2) geographical location (either on plain or in mountain). Other socio-economic characteristics X which could influence the outcomes include the owner's age, gender and education level.

4.2 Identification strategy

The OLS estimation of equation 2 is naive since two types of spurious correlation within the reference group may exist in our setting. The first type of correlation is specific to the group environment. For instance, plots within the same group share the same water source. The agricultural outcomes of each plot i are thus interdependent due to inputs contamination (e.g. fertilizers). The second type of correlation is specific to individual characteristics. For instance, farmers of the same group may help each other which depends on individual unobservable characteristics. To purge these spurious correlation other than social learning, we employ a classical IV strategy and an alternative dependent variable strategy.

explaining both the performance of the plot i in t and the group formation (and thus the performance of the group in $t-1$) may still present.

An IV strategy The first strategy relies on the random shocks of rats and pests attacks to the plots. In the context of smallholder farming, these ecological shocks cause significant damages to the rice production. As observed in Sancha village, the attacks of rats and pests are arbitrary in the time as well as in place. This observation leads to an assumption that the ecological shocks to a plot will influence its outcomes (i.e. yield and efficiency) but not that of its peers directly. On basis of this assumption, we can make use of peers' ecological shocks (i.e. average attacks of rats and pests to peers' plots) as instruments for peers' outcomes conditional on individual ecological shocks. The *exclusion restriction* of instruments will be checked by the Sargan over-identification test.

Equation 2 is thus firstly estimated using the IV estimator while controlling for group fixed effects ζ_G as follows:

$$y_{i,t} = \alpha\iota + \beta G y_{j,t-1} + \theta y_{i,t-1} + \gamma X_{i,t} + \delta G X_{j,t} + \zeta_G + \tau_t + \epsilon_{i,t}, \tag{3}$$

Still, the concern about the farmers' altruism remains. Suppose farmers help each other to mitigate the damages of ecological shocks, this altruism could improve the outcome of the group as a whole. However, this collective improvement of outcome is not due to social learning and thus bias the estimation. This source of bias is related to the unobservable individual characteristics which could be controlled for in a Within model. Therefore, to deal with individual specific correlation and validate the IV estimation, equation 2 is estimated using the Within-2SLS estimator and controlling for plot fixed effect ν_i as follows:

$$y_{i,t} = \alpha\iota + \beta G y_{j,t-1} + \theta y_{i,t-1} + \gamma X_{i,t} + \delta G X_{j,t} + \nu_i + \tau_t + \epsilon_{i,t}, \tag{4}$$

The first step regression is as follows:

$$G y_{j,t-1} = \rho_1 Grats_{j,t-1} + \rho_2 Gpests_{j,t-1} + \rho_3 \theta y_{i,t-1} + \rho_4 X_{i,t} + \rho_5 G X_{j,t} + \nu_i + \tau_t + \xi_{i,t}, \tag{5}$$

where ρ_1 and ρ_2 are the coefficients associated to the two instruments.

Alternative measure of outcome The second strategy relies on the use of technical efficiency (TE) and environmental efficiency (EE) as outcome measure. Here the TE represents farmer's managerial skill to maximize output at a given inputs level, whilst the EE represents his managerial skill to minimize environmentally detrimental input at a given output level. The use of efficiency term is of particular interest. Firstly, the managerial skill is more relevant and accurate for the social learning process. Secondly, the efficiency term is estimated from a production function. This process can purge correlations of outcomes due to contamination of productive inputs, e.g. fertilizers used in peers' plots. Thirdly, using two measures of efficiency, we can derive precise understanding about farmers willingness to learn with respect to economic performance and environment protection.

In practice, the calculation of efficiency is made by a SFA approach(Guo and Marchand, 2012)[9]. Here a farmer is assumed to use traditional inputs X and environmentally detrimental inputs N to produce rice Y on his plot i . This can be written to represent a particular technology: $Y_i = f(x_i)$, where $f(x_i)$ is a production frontier. Under the hypothesis of market imperfections, not all farmers are able to produce at the frontier. The TE thus measure the distance of Y_i from the frontier. Therefore, TE follows an output–oriented measure of production inefficiency (more conventional output with the same set of inputs). Precisely, we model the production frontier using a transcendental logarithmic ("translog") specification (Diewert, 1971) with three traditional inputs (labor, capital and water) and one environmental detrimental input of pure nitrogen (N) as follows:

$$\begin{aligned} ln(Y_{i,t}) = \beta_0 + \sum_{j=1}^{3} \beta_j ln(X_{ij,t}) + \beta_z ln(N_{i,t}) + \frac{1}{2} \sum_{j=1}^{3} \sum_{k=1}^{3} \beta_{jk} ln(X_{ji,t}) ln(X_{ki,t}) \\ + \frac{1}{2} \sum_{j=1}^{3} \beta_{jz} ln(X_{ji,t}) ln(N_{i,t}) + \frac{1}{2} \beta_{zz} ln(N_{i,t})^2 - U_{i,t} + V_{i,t}, \end{aligned} \tag{6}$$

where $i = 1, \ldots, n$ are the plots and $t = 1, \ldots, 5$ are the number of seasons; $j, k = 1, 2, \ldots, 3$ are the applied traditional inputs; $ln(Y_{i,t})$ is the logarithm of the output of

[9] See Guo and Marchand (2012) in which the authors of this present study develop a stochastic production frontier model to estimate both TE and EE. In this paper, the authors investigates the role of organic farming on EE in the same case study than the one used in this present study.

farmer i; $ln(X_{ij,t})$ is the logarithm of the j^{th} traditional input applied by the i^{th} individual; $ln(N_{i,t})$ is the logarithm of the environmental detrimental input applied by the i^{th} individual; and β_j, β_z, β_{jk}, β_{jz} and β_{zz} are parameters to be estimated[10]. $U_{i,t}$ are non-negative unobservable random variables associated with TE that follows an arbitrary distribution[11]. $V_{i,t}$ represent random shocks which are assumed to be i.i.d errors with a normal distribution of zero mean and unknown variance[12].

On the other hand, EE is defined following Reinhard et al. (1999) as the ratio of observed use of environmentally detrimental inputs (N) to the minimum feasibility, conditional to identical output and conventional inputs. In this case, EE is an input–oriented measure (less environmental detrimental input with the same output and conventional inputs) formulated by the following non–radial input–oriented measure:

$$EE_{i,t}(x,y) = [\min \theta : F(x_{i,t}, \theta Z_{i,t}) \geq y_{i,t}], \tag{7}$$

where the variable $y_{i,t}$ is the observed output for the plot i at season t, produced using x_i of the conventional inputs and N_i of the environmentally detrimental inputs. $F(.)$ is the best practise frontier with x and Z.

To save place, we do not develop the stochastic production frontier model here. A complete model is presented in our previous paper Guo and Marchand (2012) and more detailed calculation of EE can be found in Appendix A.

4.3 Technological heterogeneities

One important feature of smallholder farming is the technological heterogeneity. Specifically in our case, as farmers practice organic and conventional farming in the same group, the heterogeneity of technology could plague farmers' social learning and should be taken

[10]Similarity conditions are imposed, i.e., $\beta_{jk} = \beta_{kj}$. Moreover, the production frontier requires monotonicity (first derivatives, i.e., elasticities between 0 and 1 with respect to all inputs) and concavity (negative second derivatives). These assumptions should be checked *a posteriori* by using the estimated parameters for each data point.

[11]This can be either a half-sided normal distribution or an exponential one.

[12]A stochastic production frontier implies that the error term has two components: random shocks V_i (not attributed to the relationship between inputs and output) and an inefficiency term U_i (Aigner et al., 1977; Meeusen and van den Broeck, 1977).

account in our analysis (Munshi, 2004). To this end, we divide our sample into two sub–samples according to production technology (i.e. organic farming and conventional farming) and identify the social learning effect within each sub-sample. The hypothesis is that social learning is more likely to happen with a homogenous technology than heterogenous technology.

This approach is useful to test the constraint of technological heterogeneity for social learning. However, it also raises a *self-selection* problem. Put another way, conditions for a farmer to practice organic farming may be different from his conventional counterpart. Farmers may thus *self-select* to the organic farming. The artificial division of sample will create biased estimation if the *self-selection* exists. To rule out this potential problem, we implement a Heckman correction to the estimation (Maddala, 1983).

To do so, we estimate the probability of farmer to practice organic farming as follows:

$$Organic_{i,t} = \gamma_0 + \gamma_1 Distance_{i,t} + \gamma_2 Pollution_{i,t} + \gamma_3 y_{i,t-1} + \gamma_4 X_{i,t} + \gamma_5 GX_{j,t} + \tau_t + \varepsilon_{i,t}, \quad (8)$$

where $Organic_{i,t}$ is a dummy variable indicating the organic status of the plot i at season t. The variables $Distance_{i,t}$ is the distance for the household to reach the plot from his house. $Pollution$ is a dummy variable coding 1 if the plot has at least one neighbor using chemicals fertilizer. We assume that these two variables are exogenous factors that determine farmer's choice for organic farming (more discussion about these instruments is found in our previous paper (Guo and Marchand, 2012)). $y_{i,t-1}$ is the lagged performance of plot i. X is a matrix of plot owner's characteristics (i.e. age, gender and education level) and GX is a matrix of the same characteristics at group level. τ_t is a dummy fixing one of the five seasons. $\varepsilon_{i,t}$ is the error term.

From equation 8, we calculate the *Inverse Mills Ratio* (λ) and use it as a new control variable in the regression within each sub–sample, i.e., conventional and organic farming, as follows:

$$y_{i,t} = \alpha\iota + \beta G y_{j,t-1} + \theta y_{i,t-1} + \gamma X_{i,t} + \delta GX_{j,t} + \lambda_{i,t} + \tau_t + \epsilon_{i,t}, \quad (9)$$

In equation 9, $y_{i,t}$ is the outcome of plot i in season t, i.e. the yield, the technical efficiency and the environmental efficiency. The equation 9 is then estimated by the OLS estimator, the 2SLS estimator and the Within-2SLS estimator.

5 Data and descriptive statistics

The data used for the empirical test of social learning is derived from a household survey that we conducted in Sancha village in 2010. In collaboration with local agronomist of PCD, we have inspected 108 households that effectively participate in paddy rice production in the village. For each household, we inspected all of his paddy fields (both conventional and organic) and randomly selected two plots for the survey[13]. With the household head, we identified the selected plots and recorded information about the plot location. We then asked household head to recall information for five agricultural seasons from 2008 to 2010 with respect to organic practice, production techniques, output and inputs on each of two plots[14]. Also, a number of household characteristics including household head's age, gender and education level were recorded. Table 1 gives descriptive statistics of the main variables for this study and a summary of variable definitions can be found in Table B1 in appendix.

[13]In principle, we selected one conventional plot and one organic plot for comparison. In case of non conversion or total conversion, we selected two plots randomly.

[14]For more details, see our previous paper (Guo and Marchand, 2012)

Table 1: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Rice yield(kg/mu)	342.12	94.59	43.75	750	1,007
Labor(person h/mu)	129.82	54.12	28.45	338.81	1,007
N(kg/mu)	14.13	3.97	4.98	34.5	1,007
Capital(Yuan/mu)	73.77	52.03	3.75	265	1,007
Water (1-3)	2.41	0.61	1	3	1,007
Technical efficiency (0-1)	0.73	0.12	0.33	0.98	1,007
Environmental efficiency (0-1)	0.45	0.19	0.08	0.96	1,007
Age	54.62	12.61	28	79	1,007
Sex (1 = female)	0.6	0.49	0	1	1,007
Education (0-12 years)	3.63	3.31	0	12	1,007
Organic (1 = organic)	0.34	0.47	0	1	1,007
Distance (categorical variable)	1.92	0.87	1	4	1,007
Chemical pollution(1 = yes)	0.74	0.44	0	1	1,007
Rats (0-2)	0.49	0.57	0	2	1,007
Insects (0-2)	1.00	0.69	0	2	1,007

Source: authors' survey

From the descriptive statistics, we note that the rice production in Sancha village is principally implemented on small pieces of land (0.03 ha). It is conducted by senior (55 years) and female (60 percent) farmers, which is in line with the reality of labor outflow in the countryside. The large variation of inputs suggests that the production technology is heterogenous. For instance, about 34% of surveyed fields are under organic management, whilst 66% are under conventional farming. It is worth noting that most rice farmers have suffered ecological shocks, i.e. the attacks of pests and rats. More than half of farmers have reported to receive pests and rats attacks, so that the influence of ecological shocks should not be ignored in our analysis. Finally, we also note that the mean technical efficiency is higher (0.73) than the mean environmental efficiency (0.45) in the sample, which suggests that farmers have greater economic performance than environmental performance. In the following analysis, we will explore the dataset to investigate whether smallholder farmers could improve their performance in a social learning process and approve this participatory approach for sustainable agricultural development.

6 Econometric results

In this section, we present the identification results of social learning effect as discussed in previous section. A number of issues and policy implications raised by the results will be discussed. First of all, we regress the individual rice yield on the group rice yield to test the social learning effect as benchmark. To deal with potential group correlated effects (e.g., environmental correlation) and individual correlated effects (e.g., altruism effect), we implement Within and Within-2SLS estimation in addition to the naive OLS estimation and make a step–by–step analysis (for completeness, see Table C1 in Appendix C for first stage IV regressions). As one can note in Table 2, after correcting for group and individual correlated effects (columns 3 and 4), we do detect a significantly negative correlation of yields within the reference group, which however disappears in the Within-2SLS estimation (columns 5).

Table 2: Social learning effect and rice yield in the total sample

Dependent variable: agricultural yield					
	(1)	(2)	(3)	(4)	(5)
Lagged peers' outcome	0.066 (0.139)	-.103 (0.128)	-.807*** (0.171)	-.411** (0.16)	0.324 (0.973)
Lagged individual outcome		0.625*** (0.038)	0.592*** (0.039)	-.283*** (0.061)	-.298*** (0.081)
Age	-.344 (0.548)	-.052 (0.414)	-.405 (0.505)	-405.360*** (128.174)	-62.362 (451.420)
Sex	-17.126 (14.682)	-15.395 (11.348)	-18.599 (14.557)		
Education	1.324 (2.144)	0.037 (1.742)	1.231 (1.906)		
Rat					-13.410 (16.271)
Pests					1.571 (9.297)
Group age	7.606* (4.223)	3.253 (3.701)	-3.664 (6.208)	416.210*** (123.072)	164.542 (364.204)
Group sex	36.031 (66.600)	43.313 (64.571)	-64.025 (242.264)	-1634.042 (2155.999)	116.241 (3115.265)
Group education	-11.587 (7.048)	-7.286 (5.803)	-1.157 (7.925)	-9.679 (12.768)	-9.498 (12.262)
Intercept	229.770 (271.292)	101.094 (231.007)	1107.963** (527.650)	1579.653 (1724.860)	
Control for Lagged performance		x	x	x	x
Group dummies			x	x	x
Individual fixed effect				x	x
Instrumentation					x
Observations	805	805	805	805	805
Number of plots	202	202	202	202	202
F statistic	6.526	38.1	38.201	24.86	13.275
R^2	0.073	0.453	0.477	0.223	0.196
RMSE	183.149	140.774	138.331	97.915	114.443
Hansen statistics					1.583
P-value Hansen statistics					0.208

Estimation method: OLS estimator in columns 1 to 3, within estimator in column 4 and within-2SLS in column 5. The dependent variable is the yield defined as the raw rice output per land area. Seasons dummies are controlled for in all regressions. Group fixed effects are controlled for in column 3 and plot fixed effects in columns 4 and 5. The variables sex and education are time invariant and thus dropped in columns 4 and 5. *** statistical significance at 1%, ** statistical significance at 5%, * statistical significance at 10%.

The non-robust negative correlation is obviously not due to social learning, but rather due to contamination or spillover of inputs use (i.e. fertilizer and water). In fact, the competition of conventional and organic plots exists in the mixed total sample. It is plausible that the yield of organic plot could be negatively influenced by the overuse and leaching of chemical fertilizers from neighbor conventional plots. In turn, the yield of conventional plot is determined by chemical fertilizers which is constrained by neighbor organic plots due to conflicts and social pressure. The complex correlations of yields make it complicate to identify the social learning effect. To get rid of this concern, we now focus on the technical efficiency and environmental efficiency which are more relevant to social learning process.

Table 3 present the results when technical efficiency (columns 1 to 3) and environmental efficiency (columns 4 to 6) are used to measure plot's outcome. Columns 1 and 4 are estimated by simple OLS estimator. Columns 2 and 5 take care of environmental correlation by controlling for group dummies and applying the IV estimation. Columns 3 and 6 eliminate the altruism effect and give a causal effect of social learning by applying Within-2SLS estimation[15].

Note here in the case of technical efficiency, we detect a positive correlation among farmers when using a naive OLS estimator. However, the correlation becomes non-significant once confounding environmental factors and individual characteristics are controlled for. This result suggests the importance of group and individual correlated effects in our case, which need to be taken into account. Similarly in the case of environmental efficiency, the correlation is positive but non-significant. In contrast, the lagged individual performance is strongly significant, which suggests farmers' performance rely on their own experience, rather than others' performance. One plausible explanation to the absence of social learning effect is that in a heterogeneous population, it is difficult for farmers to learn from peers with different technology (Ellison and Fudenberg, 1993). The mixture of technology in total sample may thus plague the social learning process. To test this hypothesis, we divide the total sample into 2 sub–samples according to the production technology (i.e. organic or conventional farming). By doing so, we consider only peers who practice the same technology in the same group as learning reference.

[15]The first stage regressions can be found in Table C1 in Appendix C.

Table 3: Social learning effect and efficiency in the total sample

Dependent variable:	Technical efficiency			Environmental efficiency		
	(1)	(2)	(3)	(4)	(5)	(6)
Lagged peers' efficiency	0.003***	-.005	-.0005	0.006	0.114	1.072
	(0.0008)	(0.013)	(0.0004)	(0.023)	(2.142)	(1.006)
Lagged individual efficiency	0.978***	0.979***	0.994***	0.982***	0.975***	-.367***
	(0.0002)	(0.0009)	(0.002)	(0.006)	(0.082)	(0.063)
Age	-5.30e-07	-2.00e-06	-.0001***	-3.52e-06	0.00007	0.044
	(1.31e-06)	(1.55e-06)	(0.00004)	(0.0001)	(0.0008)	(0.03)
Sex	-7.08e-06	0.00003		0.004	0.005	
	(0.00003)	(0.0001)		(0.003)	(0.011)	
Education	9.50e-06*	1.78e-06		0.0003	0.0004	
	(5.11e-06)	(5.60e-06)		(0.0005)	(0.0005)	
Rats		-.0004**	-.00002***		0.001	0.01***
		(0.0002)	(6.15e-06)		(0.004)	(0.003)
Pests		-.00007	-7.95e-06**		-.006**	-.0002
		(0.00006)	(3.46e-06)		(0.003)	(0.002)
Group age	-.00002	-.00004***	9.68e-06	0.0002	0.001	-.052**
	(1.00e-05)	(1.00e-05)	(0.00003)	(0.0007)	(0.015)	(0.027)
Group sex	-.0007***	-.0007	0.0001	0.009	0.029	1.280**
	(0.0002)	(0.001)	(0.0006)	(0.01)	(0.157)	(0.558)
Group education	0.00005***	-1.00e-05	-2.79e-06	0.0005	0.001	0.001
	(1.00e-05)	(0.00002)	(5.72e-06)	(0.002)	(0.009)	(0.003)
Intercept	0.023***	0.03***		0.002	-.146	
	(0.0008)	(0.008)		(0.043)	(2.220)	
Group dummies		x			x	
Individual fixed effect			x			x
Instrumentation		x	x		x	x
Observations	805	805	805	805	805	805
Number of plots	202	202	202	202	202	202
F statistic	2,837,094	2,807,083	3,499,681	3,319.524	1,854.209	23.57
R^2	1	1	1	0.967	0.967	0.276
RMSE	0.0004	0.0005	0.00004	0.034	0.034	0.023
Hansen statistics		0.008	0.823		0.986	0.018
P-value Hansen statistics		0.928	0.364		0.321	0.893

Estimation method: OLS estimator in columns 1 and 4, 2LS estimator in columns 2 and 5, and within-2SLS in columns 3 and 6. The dependent variable is the estimated technical efficiency in columns 1-3, and the calculated environmental efficiency in columns 4-6. Seasons dummies are controlled for in all regressions, group fixed effects are controlled for in columns 2 and 5, and plot fixed effects in columns 3 and 6. The variables sex, education are time invariant and thus dropped in columns 3 and 6. *** statistical significance at 1%, ** statistical significance at 5%, * statistical significance at 10%.

Table 4 reports results of the conventional farming sub–sample. The estimation results are regrouped by measure of outcome, i.e. technical efficiency in columns 1 to 3, and environmental efficiency in columns 4 to 6. For each measure, we run three regressions (OLS, 2SLS and Within-2SLS) as in previous analysis[16].

As indicated in Table 4, for farmers who practice conventional farming, their performances in terms of technical and environmental efficiency, depend essentially on their own experience rather than peers' performance. This is not surprising as conventional farming is practiced in the village since long time. Without effective organization and monitoring, smallholder farmers have to handle the technology by themselves. Therefore, after a long period of self-experimentation, conventional farmers have achieved stable performance on basis of their particular conditions. Curiously, we note here female farmers have generally

[16]For more completeness, Tables C2 and C3 in Appendix C gives the first stage regressions for Tables 4 and 5 respectively.

higher economic performance (i.e. technical efficiency) than male farmers at both individual and group level. This result suggests that organization of female group is an effective way to foster farmers economic performance. This finding is interesting and important for policy design in rural areas. We will compare this result with the results of organic sub–sample.

Table 4: Social learning of conventional farming

Dependent variable:	Technical efficiency			Environmental efficiency		
	(1)	(2)	(3)	(4)	(5)	(6)
Lagged peers' efficiency	0.0007	0.011	-.0008	-.002	-.973	-.052
	(0.002)	(0.009)	(0.0006)	(0.087)	(0.63)	(0.264)
Lagged individual efficiency	0.978***	0.978***	0.993***	0.947***	0.925***	-.425***
	(0.0003)	(0.0005)	(0.001)	(0.012)	(0.018)	(0.078)
Age	4.11e-06**	5.07e-06**	-8.62e-05***	0.00009	-.00002	0.004
	(1.87e-06)	(2.08e-06)	(1.00e-05)	(0.0001)	(0.0002)	(0.008)
Sex	0.0002***	0.0002***		0.006*	0.005	
	(0.00004)	(0.00004)		(0.003)	(0.003)	
Education	7.24e-06	-1.03e-07		0.0008	0.001**	
	(6.01e-06)	(8.15e-06)		(0.0006)	(0.0007)	
Rats	-.00008*	-.00007	-2.24e-05***	0.003	0.002	0.004
	(0.00005)	(0.00005)	(8.23e-06)	(0.003)	(0.003)	(0.004)
Pests	0.00004	0.00003	-7.83e-06	-.011***	-.009***	-.004
	(0.00003)	(0.00004)	(5.15e-06)	(0.003)	(0.003)	(0.002)
Group age	0.00008***	0.00009***	-1.14e-06	0.002	0.00005	0.0008
	(0.00002)	(0.00002)	(6.36e-06)	(0.001)	(0.002)	(0.003)
Group sex	0.003***	0.003***	0.0002**	0.045***	0.008	0.014
	(0.0004)	(0.0005)	(0.00007)	(0.014)	(0.029)	(0.034)
Group education	0.0001***	7.53e-06	-3.35e-06	0.005*	0.016**	0.003
	(0.00003)	(0.0001)	(1.00e-05)	(0.003)	(0.008)	(0.004)
Inverse Mills ratio	0.00006	0.00003	-7.70e-06	0.0007	0.003	-.010
	(0.00007)	(0.00008)	(0.00002)	(0.006)	(0.006)	(0.017)
Intercept	0.016***	0.009		-.099	0.531	
	(0.002)	(0.007)		(0.08)	(0.431)	
Group dummies	x	x		x	x	
Individual fixed effect			x			x
Instrumentation		x	x		x	x
Observations	510	510	476	510	510	476
Number of plots	158	158	124	158	158	124
F statistic	2,452,183	2,290,952	2,170,216	1,345.496	1,062.598	16.176
R^2	1	1	1	0.971	0.969	0.328
RMSE	0.0004	0.0004	0.00004	0.033	0.033	0.021
Hansen statistics		0.699	0.975		0.536	0.227
P-value Hansen statistics		0.403	0.323		0.464	0.634

Estimation method: OLS estimator in columns 1 and 4, 2SLS in columns 2 and 5, and within-2SLS in columns 3 and 6. The dependent variable is the estimated technical efficiency in columns 1-3, and the calculated environmental efficiency in columns 4-6. Seasons dummies are controlled for in all regressions, group fixed effects in columns 2 and 5, and plot fixed effects in columns 3 and 6. The variables sex, education are time invariant and thus dropped in all regressions. *** statistical significance at 1%, ** statistical significance at 5%, * statistical significance at 10%.

The same estimations (OLS, 2SLS and Within-2SLS) are applied to the organic sub–sample and table 5 reports the results. For the case of organic farming which is organized in the project framework, the results are more interesting. The effect of social learning is positive for technical efficiency. The magnitude is small(i.e., 0.0001-0.0007) but significant at 1%. The result suggests that for a new technology like organic farming, farmers learn from their peers as well as their own experience. The well organized participatory social learning is thus effective and efficient to foster smallholder farmers' economic performance of organic farming.

Table 5: Social learning of organic farming

Dependent variable:	Technical efficiency			Environmental efficiency		
	(1)	(2)	(3)	(4)	(5)	(6)
Lagged peers' efficiency	0.0006**	0.0007**	0.0001***	-.0002	0.053	-.051
	(0.0003)	(0.0004)	(0.00005)	(0.041)	(0.036)	(0.041)
Lagged individual efficiency	0.977***	0.977***	0.992***	0.981***	0.981***	-.403***
	(0.0003)	(0.0002)	(0.002)	(0.013)	(0.012)	(0.11)
Age	9.35e-09	-1.77e-08	-8.17e-05***	-.0001	-.0001	0.009**
	(1.54e-06)	(1.49e-06)	(1.00e-05)	(0.0002)	(0.0002)	(0.004)
Sex	-.0001**	-.0001**		0.0001	6.63e-07	
	(0.00004)	(0.00004)		(0.005)	(0.004)	
Education	-7.22e-06	-7.38e-06		0.0007	0.0005	
	(7.82e-06)	(7.61e-06)		(0.0008)	(0.0008)	
Rats	-.0004***	-.0004***	-.00002*	0.009*	0.008*	0.018**
	(0.00007)	(0.00007)	(9.66e-06)	(0.005)	(0.005)	(0.008)
Pests	-.00007	-.00007	-1e-05**	0.004	0.003	0.005**
	(0.00004)	(0.00004)	(4.60e-06)	(0.003)	(0.003)	(0.003)
Group age	-2.09e-06	-2.76e-06	2.67e-07	-.001	-.002**	-.001*
	(7.89e-06)	(8.02e-06)	(8.69e-07)	(0.0008)	(0.0008)	(0.0006)
Group sex	0.001***	0.001***	0.00009**	0.044***	0.045***	0.033**
	(0.0003)	(0.0002)	(0.00004)	(0.016)	(0.015)	(0.016)
Group education	-9.18e-06	-1.00e-05	-.00002**	0.012	0.01	0.003
	(0.00008)	(0.00008)	(8.11e-06)	(0.008)	(0.008)	(0.005)
Inverse Mills ratio	0.00008**	0.00008**	1.00e-05	0.008**	0.009**	0.011**
	(0.00004)	(0.00004)	(8.69e-06)	(0.004)	(0.004)	(0.004)
Intercept	0.023***	0.023***		0.01	0.0003	
	(0.0003)	(0.0003)		(0.024)	(0.028)	
Group dummies	x	x		x	x	
Individual fixed effect			x			x
Instrumentation		x	x		x	x
Observations	295	295	277	295	295	277
Number of plots	97	97	79	97	97	79
F statistic	2,349,133	2,335,665	1,111,516	743.114	742.532	7.243
R^2	1	1	1	0.966	0.966	0.341
RMSE	0.0003	0.0003	0.00004	0.036	0.035	0.022
Hansen statistics		0.007	0.174		0.118	1.342
P-value Hansen statistics		0.935	0.677		0.731	0.247

Estimation method: OLS estimator in columns 1 and 4, 2SLS in columns 2 and 5, and within-2SLS in columns 3 and 6. The dependent variable is the estimated technical efficiency in columns 1-3, and the calculated environmental efficiency in columns 4-6. Seasons dummies are controlled for in all regressions, group fixed effects in columns 2 and 5, and plot fixed effects in columns 3 and 6. The variables sex, education are time invariant and thus dropped in all regressions. *** statistical significance at 1%, ** statistical significance at 5%, * statistical significance at 10%.

Nevertheless, the social learning effect is non-significant for environmental efficiency. Put differently, smallholder farmers are economic rational rather than environmental protectionist. They learn to maximize the yield but not to minimize the use of environmentally detrimental nitrogen, even in the case of organic farming. This result indicates the critical limitation of social learning process in promoting resource conservative agriculture. Without social learning, farmers will take long time to improve their environmental efficiency based on their own experience and external incentives. Therefore, government's guidance and assistance become necessary in the urgent situation of environment deterioration. In the case of Sancha village, more environmental education and extension service should be provided to support smallholder farmers.

To check the effect of female groups, we find that the role of female groups is even more important in the case of organic farming. Farmers in female groups have significantly higher performances in terms of both technical efficiency and environmental efficiency. This result is in line with our previous finding that women favor the adoption of organic farming in Sancha village (Renard and Guo, 2013). This evidence has direct policy implication for the human development in rural China. In a time where women are becoming major labor force in agriculture (De Brauw et al., 2013), policy design should be more favorable with respect to the education and organization of women in rural areas. In any circumstance, women's interests, specificity and ability should be well recognized to promote a more performing and resource conserving agriculture.

7 Discussion and conclusion

As a prevalent form of agriculture, smallholder farming plays a crucial role in the sustainable agricultural development in developing countries. In order to empower smallholder farmers for ecological innovation such as organic farming, initiatives of participatory social learning are put forward and experienced within the framework of New Rural Reconstruction in China. The hypothesis is that smallholder farmers could learn from each other to revive local knowledge and improve their performance. If it is true, farmers could rely on themselves and reduce their dependance on external assistance for sustainable agricultural development.

This paper aims to test this hypothesis with the experience in Sancha village from southwest China. We estimate the social learning effect on smallholder farmers' performances in a Spatial Autoregressive (SAR) model. To disentangle the social learning effect from environment related contamination effect and other individual related altruism effect, we make use of ecological shocks (i.e., rats and pests attacks) as instruments to run a

IV-2SLS estimation and use technical efficiency and environmental efficiency as dependant variables. To investigate the technological constraints for social learning, we separate the sample by farmers' technology (i.e., conventional and organic farming) and compare the identification results.

In a step by step analysis, we demonstrate that social learning is conditional on the same technology. Precisely, farmers mainly depend on their own experience in the case of conventional farming. In the case of organic farming, the social learning effect is significant. Farmers learn from their peers as well as their own experience. This result suggests that for new technology such as organic farming, the organization of farmers for participatory social learning is an effective and efficient way to adapt the technology to local conditions and improve smallholder farmers' performance.

However, social learning is not an excuse to withdraw the agricultural extension service. We detect that smallholder farmers are economic rational rather than environment protectionist. Given their poor economic condition, smallholder farmers are more interested by the improvement of technical efficiency rather than the environmental efficiency. Without effective monitoring, the goal of environment protection will not be achieved through a social learning process, even in the case of organic farming.

By recognizing the limitation of social learning, our policy recommendation stresses the revival of agricultural extension system in rural areas. Provided the public finance constraints, the extension system has almost collapsed in China (Huang et al., 2004; Jin et al., 2009). In absence of agricultural extension service, farmers are driven by economic interests and pursue the only objective of agricultural productivity. The serious consequence has been illustrated by increasing use of chemical inputs and environmental deterioration. Our study confirms once again the risk due to absence of effective agricultural extension service. Even using a environmental friendly technology such as organic farming, smallholder farmers' economic incentives will never change.

Finally, we recommend more attention to women in the sustainable agricultural development. As suggested by our study, female groups favor the smallholder performance, especially in the case of organic farming. This is probably related to women's advantage in communication, sensibility and availability, which is favorable for sustainable agriculture. Therefore, we should provide more opportunity and resource to educate and organize women in the design of sustainable agricultural program. We believe that the human and social capital in a feminized agriculture are the real assets for sustainable agricultural development in China.

References

Aigner, D., Lovell, C. A. K., Schmidt, P., 1977. Formulation and estimation of stochastic frontier production function models. Journal of Econometrics 6 (1), 21–37.

Altieri, M. A., 2002. Agroecology: the science of natural resource management for poor farmers in marginal environments. Agriculture, ecosystems & environment 93 (1), 1–24.

Bandiera, O., Rasul, I., 2006. Social networks and technology adoption in northern mozambique. The Economic Journal 116 (514), 869–902.

Bandura, A., 1977. Social learning theory. Journal of the International Differential Treatment Association.

Battese, G. E., Coelli, T. J., 1992. Frontier production functions, technical efficiency and panel data: With application to paddy farmers in India. Journal of Productivity Analysis 3 (1-2), 153–169.

Bélières, J., Bonnal, P.and Bosc, P. L. B. M. J., Sourisseau, J., 2013. Les agricultures familiales du monde. définitions, contributions et politiques publiques.

Besley, T., Case, A., 1994. Diffusion as a learning process: evidence from hyv cotton.

Braun, A. R., Thiele, G., Fernández, M., 2000. Farmer field schools and local agricultural research committees: complementary platforms for integrated decision-making in sustainable agriculture. ODI London.

Brock, W. A., Durlauf, S. N., 2001. Interactions-based models. Handbook of econometrics 5, 3297–3380.

Cai, F., Wang, M., 2008. A counterfactual analysis on unlimited surplus labor in rural china. China and World Economy 16 (1), 51?5.

Conley, T., Christopher, U., 2001. Social learning through networks: The adoption of new agricultural technologies in ghana. American Journal of Agricultural Economics 83 (3), 668–673.

Conley, T. G., Udry, C. R., 2010. Learning about a new technology: Pineapple in ghana. The American Economic Review 100 (1), 35–69.

De Brauw, A., Huang, J., Zhang, L., Rozelle, S., 2013. The feminisation of agriculture with chinese characteristics. The Journal of Development Studies 49 (5), 689–704.

Desai, V., Potter, R. B., 2013. The Companion to Development Studies. Routledge.

Diewert, W. E., 1971. An application of the shephard duality theorem: A generalized leontief production function. The Journal of Political Economy 79 (3), 481–507.

Ellison, G., Fudenberg, D., August 1993. Rules of thumb for social learning. Journal of Political Economy 101 (4), 612–43.

Fan, S., Chan-Kang, C., 2005. Is small beautiful? farm size, productivity, and poverty in asian agriculture. Agricultural Economics 32, 135–146.

FAO, 2002. Organic Agriculture, Environment and Food Security. FAO.

Fazey, I., Fazey, J. A., Fischer, J., Sherren, K., Warren, J., Noss, R. F., Dovers, S. R., 2007. Adaptive capacity and learning to learn as leverage for social-ecological resilience. Frontiers in Ecology and the Environment 5 (7), 375–380.

Feder, G., 1985. The relation between farm size and farm productivity: The role of family labor, supervision and credit constraints. Journal of Development Economics 18 (2), 297–313.

Feder, G., Murgai, R., Quizon, J. B., 2004. The acquisition and diffusion of knowledge: The case of pest management training in farmer field schools, indonesia. Journal of Agricultural Economics 55 (2), 221–243.

Folke, C., Hahn, T., Olsson, P., Norberg, J., 2005. Adaptive governance of social-ecological systems. Annual Review of Environment and Resource 30, 441–473.

Foster, A. D., Rosenzweig, M. R., 1995. Learning by doing and learning from others: Human capital and technical change in agriculture. Journal of political Economy, 1176–1209.

Godtland, E. M., Sadoulet, E., De Janvry, A., Murgai, R., Ortiz, O., 2004. The impact of farmer field schools on knowledge and productivity: A study of potato farmers in the peruvian andes. Economic Development and Cultural Change 53 (1), 63–92.

Granovetter, M., 1978. Threshold models of collective behavior. American journal of sociology, 1420–1443.

Greenland, D. J., 1997. The sustainability of rice farming. CABI.

Guo, H., Marchand, S., 2012. The environmental efficiency of organic farming in developing countries: a case study from china. CERDI-Etudes et Documents 2012.38 (2012.38).

Hazell, P., Poulton, C., Wiggins, S., Dorward, A., 2007. The Future of Small Farms for Poverty Reduction and Growth. International Food Policy Research Institute.

Hazell, P. B. R., Ramasamy, C., 1991. The Green Revolution reconsidered: the impact of high-yielding rice varieties in South India. Johns Hopkins University Press.

Heltberg, R., 1998. Rural market imperfections and the farm size-productivity relationship: Evidence from pakistan. World Development 26 (10), 1807–1826.

Huang, J., Hu, R., Scott, R., 2004. China's agricultural research system and reforms: challenges and implications for developing countries. Asian Journal of Agriculture and Development 1 (1), 98–112.

IFAD, Damiani, O., 2002. The Adoption of Organic Agriculture Among Small Farmers in Latin America and the Caribbean: Thematic Evaluation. International Fund for Agricultural Development.

Jin, S., Huang, J., Rozelle, S., 2009. The production performance of china's transforming agriculture. Choices 24 (4).

Jovanovic, B., Nyarko, Y., 1994. Learning by doing and the choice of technology. Tech. rep., National Bureau of Economic Research.

Lee, L.-f., October 2007. Identification and estimation of econometric models with group interactions, contextual factors and fixed effects. Journal of Econometrics 140 (2), 333–374.

Lipton, M., 2006. Can small farmers survive, prosper, or be the key channel to cut mass poverty. Journal of Agricultural and Development Economics 3 (1), 58–5.

Lucas, R. E., 1988. On the mechanics of economic development. Journal of monetary economics 22 (1), 3–42.

Maddala, G. S., 1983. Limited-dependent and qualitative variables in econometrics. Vol. 3. Cambridge University Press.

Mahajan, V., Muller, E., Bass, F. M., 1990. New product diffusion models in marketing: A review and directions for research. The Journal of Marketing, 1–26.

Manski, C. F., 1993. Identification of endogenous social effects: The reflection problem. The review of economic studies 60 (3), 531–542.

Meeusen, W., van den Broeck, J., 1977. Efficiency estimation from cobb-douglas production functions with composed error. International Economic Review 18 (2), 435–444.

Moffitt, R. A., 2001. Policy interventions, low-level equilibria, and social interactions. Social dynamics, 45–82.

Munshi, K., 2004. Social learning in a heterogeneous population: technology diffusion in the indian green revolution. Journal of Development Economics 73 (1), 185–213.

NBSC, 2012. 2011 monitoring report of migrant workers in china. Tech. rep., National Bureau of Statistics of China.

Pretty, J., Morison, J., Hine, R., 2003. Reducing food poverty by increasing agricultural sustainability in developing countries. Agriculture, Ecosystems & Environment 95 (1), 217–234.

Pretty, J., Uphoff, N., 2002. Social and human capital for sustainable agriculture., 47–57.

Pretty, J. N., 1995. Participatory learning for sustainable agriculture. World Development 23 (8), 1247–1263.

Raghbendra, Chitkara, P., Gupta, S., 2000. Productivity, technical and allocative efficiency and farm size in wheat farming in india: a DEA approach. Applied Economics Letters 7 (1), 1–5.

Reinhard, S., Lovell, C. A. K., Thijssen, G., 1999. Econometric estimation of technical and environmental efficiency: An application to dutch dairy farms. American Journal of Agricultural Economics 81 (1), 44–60.

Renard, M.-F., Guo, H., 2013. Social activity and collective action for agricultural innovation: a case study of new rural reconstruction in china. CERDI-Etudes et Documents 2013.06 (2013.06).

Rogers, E. M., 1995. Diffusion of innovations. Simon and Schuster.

Romer, P. M., 1986. Increasing returns and long-run growth. The Journal of Political Economy, 1002–1037.

Sen, A., 1962. An aspect of indian agriculture. Economic Weekly Annual Number 14.

Stoop, W. A., Uphoff, N., Kassam, A., 2002. A review of agricultural research issues raised by the system of rice intensification (SRI) from madagascar: opportunities for improving farming systems for resource-poor farmers. Agricultural Systems 71 (3), 249–274.

Swaminathan, M. e. a., 2013. Investing in smallholder agriculture for food security.

Swinton, S. M., Lupi, F., Robertson, G. P., Hamilton, S. K., 2007. Ecosystem services and agriculture: Cultivating agricultural ecosystems for diverse benefits. Ecological Economics 64 (2), 245–252.

Twarog, S., 2006. Organic agriculture: A trade and sustainable development opportunity for developing countries. Trade and Environment Review, 141–223.

Van den Berg, H., Jiggins, J., 2007. Investing in farmers: the impacts of farmer field schools in relation to integrated pest management. World Development 35 (4), 663–686.

Wilson, R., 1975. Informational economies of scale. The Bell Journal of Economics, 184–95.

WorldBank, 2009. World Development Report 2010: Development and Climate Change. World Bank Publications.

Young, H. P., 2009. Innovation diffusion in heterogeneous populations: Contagion, social influence, and social learning. The American economic review 99 (5), 1899–1924.

Appendix A The calculation of the environmental efficiency

The logarithm of the output of a technically efficient producer $Y_{i,t}^F$ with $X_{i,t}$ and $Z_{i,t}$ can be obtained by setting $U_{i,t} = 0$ in Equation 6. However, the logarithm of the output of an environmentally efficient producer $Y_{i,t}$ with $X_{i,t}$ and $Z_{i,t}$ is obtained by replacing $Z_{i,t}$ by $Z_{i,t}^F$, where $Z_{i,t}^F = EE_{i,t} * Z_{i,t}$, and setting $U_{i,t} = 0$ in Equation 6 as follows

$$\begin{aligned} ln(Y_{i,t}) = \beta_0 + \sum_{j=1}^{3} \beta_j ln(X_{ij,t}) + \beta_z ln(Z_{i,t}^F) + \frac{1}{2}\sum_{j=1}^{3}\sum_{k=1}^{3} \beta_{jk} ln(X_{ji,t}) ln(X_{ki,t}) \\ + \frac{1}{2}\sum_{j=1}^{3} \beta_{jz} ln(X_{ji,t}) ln(Z_{i,t}^F) + \frac{1}{2}\beta_{zz} ln(Z_{i,t}^F)^2 + V_{i,t}, \end{aligned} \tag{10}$$

The logarithm of EE ($lnEE_{i,t} = lnZ_{i,t}^F - lnZ_{i,t}$) can now be calculated by setting Equations 6 and 10 equal as follows:

$$\frac{1}{2}\beta_{zz}(lnEE_{i,t})^2 + (lnEE_{i,t})[\beta_z + \sum_{j=1}^{3} \beta_{jz} lnX_{ij,t} + \beta_{zz} lnZ_{i,t}] + U_{i,t} = 0, \tag{11}$$

so,

$$\begin{aligned} & lnEE_{i,t} \\ & = \left[- \left(\overbrace{\beta_z + \sum_{j=1}^{3} \beta_{jz} lnX_{ij,t} + \beta_{zz} lnZ_{i,t}}^{A} \right) \right. \\ & \left. \pm \left\{ \left(\overbrace{\beta_z + \sum_{j=1}^{3} \beta_{jz} lnX_{ij,t} + \beta_{zz} lnZ_{i,t}}^{B} \right) - 2\beta_{zz} U_{i,t} \right\}^{0.5} \right] / \beta_{zz} \end{aligned} \tag{12}$$

As mentioned by Reinhard et al. (1999), the output-oriented efficiency is estimated econometrically whereas environmental efficiency (Eq. 11) is calculated from parameter estimates (β_z and β_{zz}) and the estimated error component ($U_{i,t}$).

As we have mentioned, a technically efficient farm ($U_{i,t} = 0$) is necessarily environmentally efficient ($lnEE_{i,t} = 0$). Thus, the "$+\sqrt{}$" must be used[17].

The final empirical model estimated in the translog case is:

$$Output_{k,i,t} = \beta_0 + \beta_1.Labor_{k,i,t} + \beta_2.Capital_{k,i,t} + \beta_3.Water_{k,i,t} + \beta_4.N_{k,i,t} + \beta_5.Labor^2_{k,i,t} + \cdots + \beta_9.Labor_{k,i,t} * Capital_{k,i,t} + \cdots + \sum_{j=1}^{5} Seasons + \sum_{j=1}^{7} SEED - U_{k,i,t} + V_{k,i,t}, \tag{13}$$

which represents the relationship between the output and both traditional and environmental inputs of plot k for farmer i and where $Seasons$ is a dummy fixing each of the five crop seasons and $SEED$ is a dummy variable for the type of species of rice. The output is the yield of raw rice harvested from the plot at end of the season. Traditional inputs are (1) the labor defined as the number of hours spent in paddy rice production on the plot weighted by the age of farmer ("hours/mu"), (2) capital defined as money spent for the rice production on the plot including the machinery, employment and seed cost ("yuan/mu"), and (3) water calculated from an index of water availability to the plot, range from 1 (weak water availability) to 3 (good water availability). The environmental input (N) is the use of pure nitrogen which is the most important nutrient input for paddy rice production and also the biggest pollutant to underground water and air resulting from agricultural production in China (see Table B1 for description and definition of variables).

Finally, the inefficiency term is allowed to be time–variant following the Battese–Coelli parametrization of time-effects (Battese and Coelli, 1992). Therefore, the maximum likelihood estimator is used to estimate TE, which is modeled as a truncated-normal random variable multiplied by a specific function of time[18].

[17]The sign in front of the term B should be necessarily positive. Thus, if $U_{i,t} = 0$, then $lnEE_{i,t} = 0$.

[18]Estimations are made using Stata 11 and the command *xtfrontier*.

Appendix B Definition of variables

Table B1: Definition of variables

Variable Name	Definition and description
Organic	Farmer's self report organic status. It's a binary variable coded "1" if the plot is under organic management, "0" otherwise.
Yield	The quantity of raw rice harvested from the plot at end of the season, the unit is "jin/mu".
Labor	Hours spent in paddy rice production on the plot. It is weighted by the age of farmer. The unit is "hours/mu".
N	The external Nitrogen input from organic source or inorganic source for the paddy rice production on the plot. The unit is "jin/mu".
Capital	Money spent for the rice production on the plot including the machinery, employment and seed cost. The unit is "yuan/mu".
Water	Index of water availability to the plot, range from 1 to 3. High index means good water availability.
Age	The age of the household head (in years).
Sex	The Sex of the household head: 1 = female.
Education	Years of education of the household head.
Distance	The geographical distance from farmer's house to the plot. Measured in minutes of walk. Range from 1 to 4.
Chemical pollution	The presence of pollution from chemical fertilizer application nearby the plot: With"1" yes and "0" no.
TE	Technical efficiency calculated from the SFA model.
EE	Environmental efficiency calculated using the method of Reinhard et al. (1999).
Rats	The damage caused by rats attacks. With "0" No damage, "1" I level damage and "2" II level damage.
Pests	The damage caused by pests attacks. With "0" No damage, "1" I level damage and "2" II level damage.

Appendix C First stage IV regressions

Table C1: First stage IV regressions in the all sample

Dependent variable:	Lag peers' Yield	Lag peers' technical efficiency		Lag peers' environmental efficiency	
First stage reg. of	Col.5-Table 2	Col.2-Table 3	Col.3-Table 3	Col.5-Table 3	Col.6-Table 3
Lag group rats (t-1)	-33.963***	0.012	0.052*	-.003	0.003**
	(8.502)	(0.018)	(0.03)	(0.003)	(0.001)
Lag group pests (t-1)	6.618	0.0005	-.006	0.0004	0.003***
	(5.702)	(0.013)	(0.014)	(0.002)	(0.001)
Lag individual outcome	0.024**	0.056***	2.040***	-.038***	-.009
	(0.012)	(0.02)	(0.533)	(0.002)	(0.007)
Age	-393.896***	-.00004	-.017	-.0004***	0.014***
	(15.234)	(0.0001)	(0.013)	(0.00009)	(0.002)
Sex		0.01***		-.005**	
		(0.003)		(0.002)	
Education		0.0001		0.00003	
		(0.0005)		(0.0001)	
Rats	5.073*	-.011***	-.0002	0.001	-.0004
	(2.961)	(0.004)	(0.003)	(0.0009)	(0.0004)
Pests	3.028	-.004*	-.002	0.0008	-.001***
	(1.953)	(0.002)	(0.002)	(0.0007)	(0.0004)
Group age	344.269***	0.0005	0.005	-.007***	0.006***
	(13.796)	(0.001)	(0.012)	(0.002)	(0.002)
Group sex	-2550.105***	0.087*	-.867***	-.072	-.371***
	(267.119)	(0.048)	(0.288)	(0.051)	(0.03)
Group education	0.477	-.001	-.003	0.004**	0.0004
	(4.113)	(0.002)	(0.003)	(0.002)	(0.0008)
Intercept	4925.216***	0.61***	0.454	0.978***	-.450***
	(264.132)	(0.115)	(0.294)	(0.142)	(0.027)
Observations	805	805	805	805	805
Number of plots	202	202	202	202	202
F statistic	904.286	53.961	31.663	4272.102	6506.258
R^2	0.766	0.396	0.234	0.974	0.938
RMSE	25.096	0.038	0.024	0.011	0.004

Note: Robust standard errors. Seasons dummies are controlled for in all regressions, group fixed effects are controlled for in columns 2 and 4, and plot fixed effects are controlled for in columns 1, 3 and 5. The variables sex, education are time invariant and thus dropped in all regressions. *** statistical significance at 1%, ** statistical significance at 5%, * statistical significance at 10%.

Table C2: First stage IV regressions and conventional farming

First stage regression of	Col.2-Table 4	Col.3-Table 4	Col.5-Table 4	Col.6-Table 4
Lag group rats	-.013**	0.011*	-.009	0.027**
	(0.006)	(0.006)	(0.01)	(0.01)
Lag group pests	-.006***	-.009***	-.009***	-.013***
	(0.002)	(0.002)	(0.002)	(0.002)
Lag individual outcome	-.033***	0.087	-.036***	-.082***
	(0.004)	(0.15)	(0.005)	(0.022)
Age	-.0001***	0.017***	-.0002***	0.032***
	(0.00004)	(0.002)	(0.00006)	(0.002)
Sex	-.001		-.003**	
	(0.0008)		(0.001)	
Education	0.0006***		0.0009***	
	(0.0002)		(0.0002)	
Rats	-.001	0.00006	-.002	-.00003
	(0.0009)	(0.001)	(0.001)	(0.002)
Pests	0.0009	-.001	0.001	-.002*
	(0.0007)	(0.0009)	(0.001)	(0.001)
Group age	-.001***	-.004***	-.004***	-.006***
	(0.0004)	(0.001)	(0.0007)	(0.001)
Group sex	-.030***	-.038***	-.040***	-.058***
	(0.005)	(0.012)	(0.008)	(0.019)
Group education	0.01***	0.001	0.014***	-.005
	(0.001)	(0.004)	(0.002)	(0.004)
Inverse Mills Ratio	0.002	-.004	0.004	-.008
	(0.002)	(0.004)	(0.002)	(0.006)
Intercept	0.874***	0.007	0.734***	-.920***
	(0.03)	(0.054)	(0.055)	(0.064)
Observations	510	510	510	510
Number of plots	158	158	158	158
F statistic	726.302	196.986	1042.97	295.084
R^2	0.967	0.741	0.968	0.831
RMSE	0.009	0.006	0.014	0.008

Note: The dependent variable is the lag of peers' technical efficiency in columns 1 and 2, and the lag of peers' environmental efficiency in columns 3 and 4. Robust standard errors. Seasons dummies and plot fixed effects are controlled for in all regressions. The variables sex, education are time invariant and thus dropped in columns 2 and 4. *** statistical significance at 1%, ** statistical significance at 5%, * statistical significance at 10%.

Table C3: First stage IV regressions and organic farming

First stage regression of	Col.2-Table 5	Col.3-Table 5	Col.5-Table 5	Col.6-Table 5
Lag group rats	-.507***	-.532***	-.329***	-.358***
	(0.07)	(0.105)	(0.031)	(0.054)
Lag group pests	-.016**	-.021*	-.034***	-.049***
	(0.006)	(0.011)	(0.009)	(0.016)
Lag individual outcome	-.017	0.853	-.012	0.182*
	(0.018)	(0.614)	(0.013)	(0.1)
Age	-.00008	-.0008	-.00003	0.01***
	(0.0002)	(0.008)	(0.0002)	(0.004)
Sex	-.003		-.006	
	(0.004)		(0.005)	
Education	0.0004		0.0005	
	(0.001)		(0.001)	
Rats	0.01**	0.022**	0.008**	0.015
	(0.004)	(0.01)	(0.004)	(0.01)
Pests	0.001	-.0002	0.003	-.004
	(0.003)	(0.003)	(0.004)	(0.003)
Group age	0.002	0.001	0.001	0.0002
	(0.002)	(0.002)	(0.002)	(0.001)
Group sex	-.026*	-.054	-.038**	-.047
	(0.015)	(0.07)	(0.018)	(0.045)
Group education	0.011	0.011	-.002	0.002
	(0.016)	(0.011)	(0.022)	(0.008)
Inverse Mills Ratio	0.0005	0.012	0.004	0.018*
	(0.005)	(0.012)	(0.007)	(0.01)
Intercept	0.834***	0.298	0.642***	0.007
	(0.114)	(0.19)	(0.071)	(0.161)
Observations	295	295	295	295
Number of plots	97	97	97	97
F statistic	4611.076	120.862	647.641	68.41
R^2	0.936	0.803	0.881	0.689
RMSE	0.032	0.027	0.034	0.025

Note: The dependent variable is the lag of peers' technical efficiency in columns 1 and 2, and the lag of peers' environmental efficiency in columns 3 and 4. Robust standard errors. Seasons dummies and plot fixed effects are controlled for in all regressions. The variables sex, education are time invariant and thus dropped in columns 2 and 4. *** statistical significance at 1%, ** statistical significance at 5%, * statistical significance at 10%.